CENDI Meeting
January 8, 2015

CODATA
CREATING THE DATA FUTURE
Discussion of National and International Data Science Organizations

Bonnie C. Carroll
Executive Committee, CODATA
bcarroll@iiaweb.com
www.codata.org/blog
@simonhodson99

CODATA History and Mission

Fredrick Rossini

NIST CODATA @ 45 Years: the story of the ICSU Committee on Data for Science and Technology

By the early **1960s** a number of scientific leaders began to realize that this deluge of data was swamping the traditional publication and retrieval mechanisms, and that there was a danger that much of it would be lost to future generations.

an organized international effort was needed to improve the management and preservation of scientific data and to facilitate coordination among interested groups throughout the world,

the creation of CODATA as an interdisciplinary Scientific Committee of the International Council for Science (ICSU), was established in 1966

to promote and encourage, on a world-wide basis, the compilation, evaluation and dissemination of reliable numerical data of importance to science and technology.

CODATA Mission

ICSU
International Council for Science

ICSU's Mission → **CODATA's Mission**

*"Strengthen international science for the benefit of society **by promoting improved scientific and technical data management and use**."*

- CODATA works to improve the quality, reliability, management and accessibility of data of importance to all fields of science and technology.
- CODATA is concerned with all types of data. Particular emphasis is given to data management problems common to different disciplines and to data used outside the field in which they were generated.

CODATA Structure

- Governed by its General Assembly and Executive Committee
- Committees
- Task Groups
- Working Groups
- National Members (Committees)

- Co-operation with other organizations on common interests (including Union activities)

Four Primary Product and Service Areas

- Biennial CODATA International Conference which attracts data specialists from around the world.
- Specialist meetings of scientific data experts, to address issues specific to one discipline or topic.
- Publications on data management, data compilation, surveys of data activities, and conference proceedings.
- Sponsorship of Task Groups, Working Groups, Commissions and other groups addressing specific data issues

CODATA Members

- **National Members:** a scientific academy, research council, scientific institution – or association of such institutions – that has activities to promote Open Scientific Data and wishes to represent a given country as CODATA member.

  23 Countries

- **Union Members:** any scientific union federated in ICSU that wishes to collaborate with CODATA's mission of promoting Open Scientific Data.

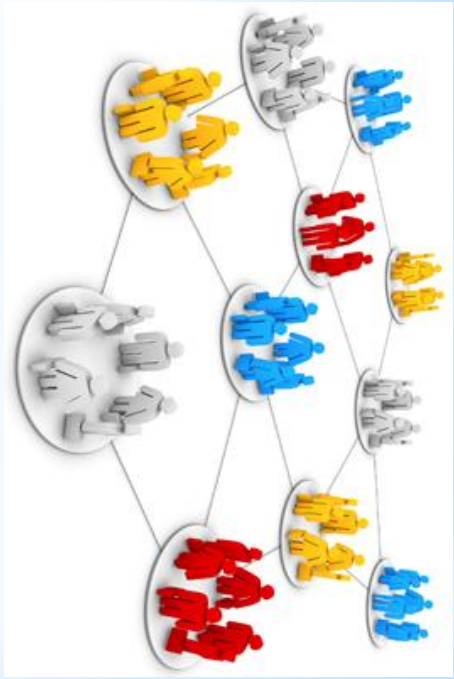
  16 International Scientific Union Members

- **Affiliate Members:** includes organizations that support and wish to collaborate with CODATA's mission of promoting Open Scientific Data.

  25 Affiliate Member Organizations

CODATA National Committees

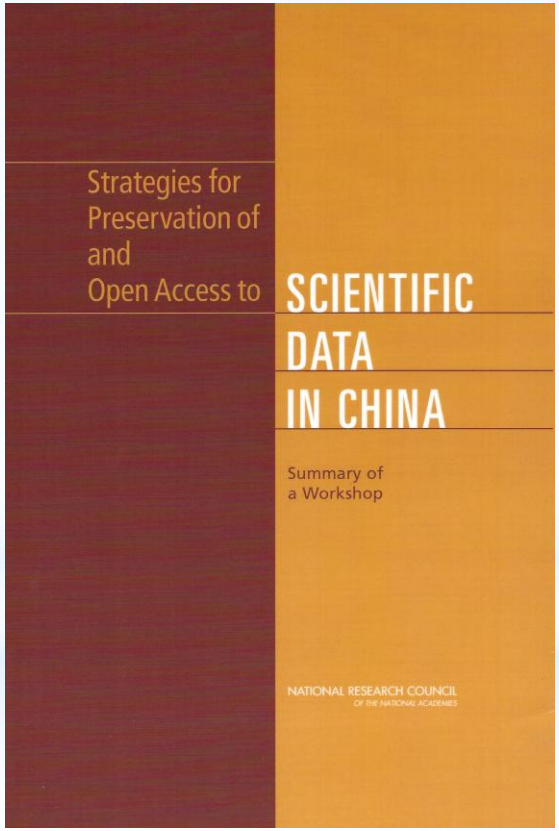
- **What are the benefits of having a CODATA National Committee?**
  - Point of contact and engagement with CODATA;
  - Group to build and coordinate collaboration on CODATA agendas and activities (including data policies and standards, Task Groups, capacity building and training, data issues in developing countries, early career data professionals);
  - Entity to collaborate with other National Committees, bilaterally or in groups;
  - Forum by which national stakeholders (research funders, National Academies, research institutions, data centers, learned societies, research libraries, etc) may raise and advance various agendas particularly those with an international dimension.



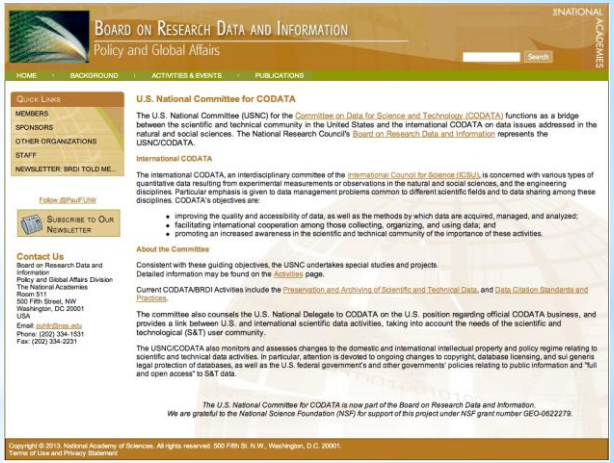
Active CODATA National Committees

CODATA US-China Workshop, 2006

CODATA China:
http://www.codata.cn/en/enindex.asp



CODATA US under auspices of BRDI NAS:
http://sites.nationalacademies.org/pga/brdi/PGA_046807



CODATA's Strategic Plan, 2013-18

Data Policy

- International and national aspects of data policy.
- Data policy committee: setting an international agenda for data policy, expert forum, advice and consultancy.
- Coordinating with national committees.

Data Science

- Long-standing activities: fundamental constants.
- Strategic working groups; community-driven task groups.
- Disciplinary and interdisciplinary data challenges, Big Data

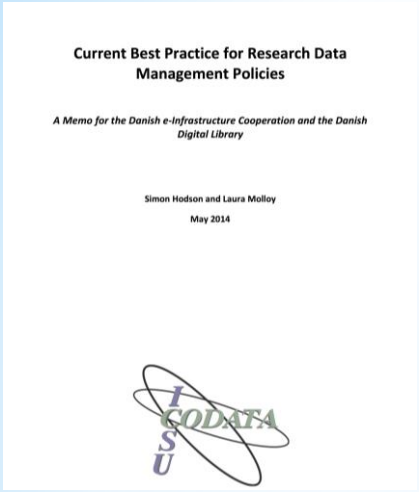
Data for International Science

- Address Data issues and challenges in international, Interdisciplinary science programs including ICSU Future Earth and Integrated Research on Disaster Risk (IRDR)
- Work closely with international partners e.g. WDS, RDA. GEO

Capacity Building

- Longstanding work on data preservation and access with developing countries.
- Executive Committee Task Force on Capacity Building: setting an international agenda for capacity building; Early Career WG

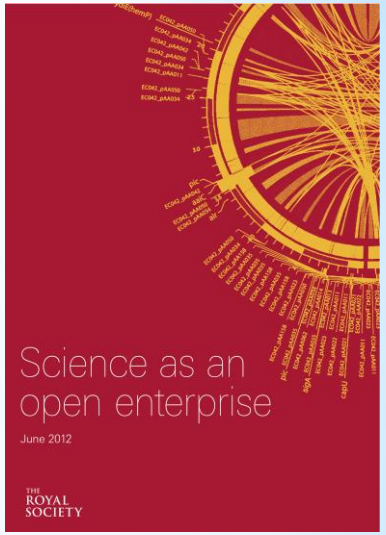
Data Policies

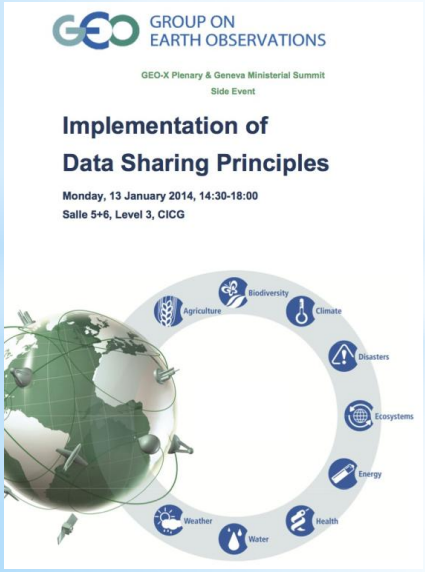


Expert Report on Data Policies for Danish e-Infrastructure Group

COUNCIL OF FINNISH ACADEMIES

Workshop on data policies for ICSU Europe planned for Spring 2015





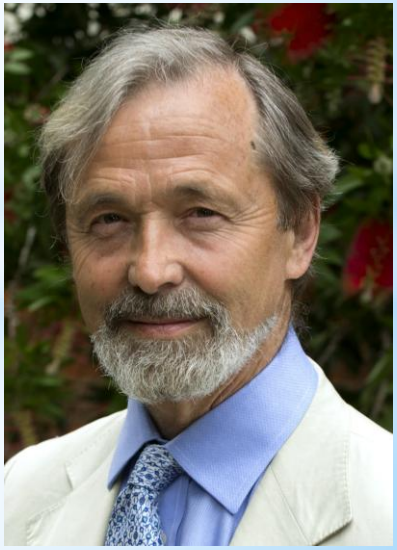
Substantial input to ICSU Report on Statement on Open Access and Metrics
http://bit.ly/icsu-OA-statement

ICSU
International Council for Science

Leading role in GEO DSWG and DMP TF
http://bit.ly/GEO_DSPs

GEO

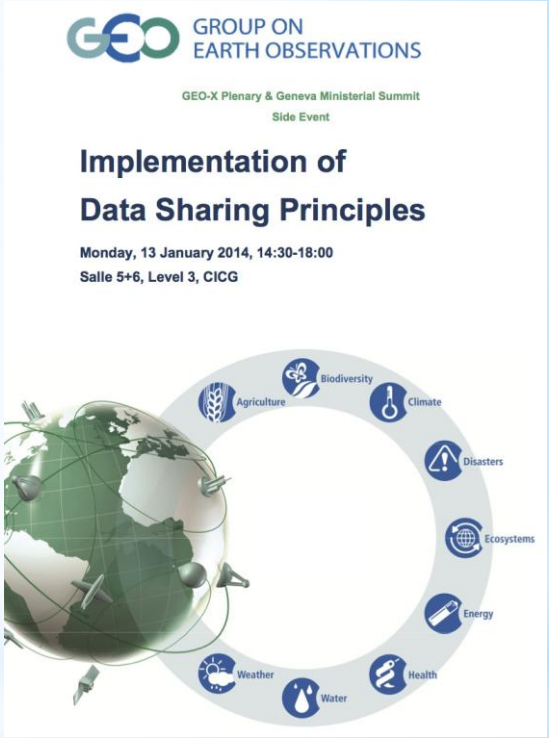
New CODATA President Geoffrey Boulton, FRS Chair of *Science as an Open Enterprise* Report



GEO Data Sharing Principles, post 2015

GROUP ON EARTH OBSERVATIONS

- Data, metadata and products will be shared through GEOSS as **Open Data by default**, by making them available as part of the GEOSS Data-CORE without charge, without restrictions on reuse, subject to the conditions of registration and attribution when the data are reused;
- Where international instruments, **national policies or legislation preclude the sharing of data** as Open Data they should be made **available through GEOSS with minimal restrictions on use and at no more than the cost of reproduction and distribution**;
- All shared data, products and metadata will be made available through GEOSS with **minimum time delay**.



GEO Data Sharing Principles, post-2015: http://bit.ly/GEO_DSPs

Data Principles for Future Earth

- Data principles proposed by CODATA and WDS.
  - **Excellence in Data Management:** data generated and modified in Future Earth, and associated research products—such as code—will be managed throughout the research lifecycle in accordance with good practice.
  - **Openness and Protection:** data, and other research products, generated and modified in Future Earth will be made as openly available as possible, with minimum delay and at minimum cost, while respecting relevant international agreements, national policies and legislation for the protection of personal, sensitive and commercial data;
  - **Integrity and Legacy:** data, and other research products, generated and modified in Future earth will be discoverable, accessible, intelligible and reusable, in the short and long term, and will therefore be selected appropriately, quality-assessed, furnished with appropriate metadata, machine readable licenses, and maintained in trusted digital repositories.

futureearth
research for global sustainability

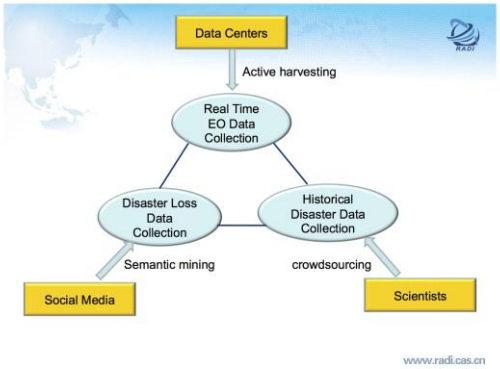
Challenges in Data Science

Advancing Informatics for Microbiology

IUMS IUBS WDCM WORLD DATA CENTRE FOR MICROORGANISMS

Collaboration around interoperability and knowledge extraction issues; research and training workshops: http://www.codata.org/task-groups/advancing-informatics-for-microbiology

Linked Open Data for Global Disaster Risk Research



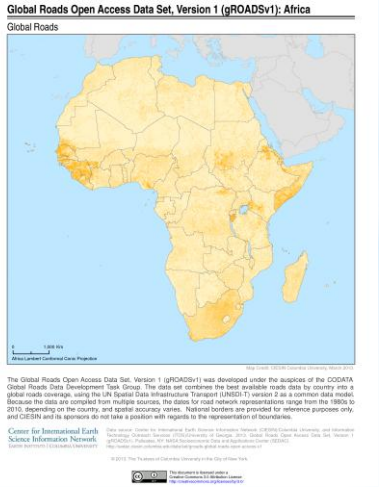
Preparing White Paper on Use of LOD for Disaster Data: http://www.codata.org/task-groups/linked-open-data-for-global-disaster-risk-research

Data Science Journal

DSJ

New platform and partnership for CODATA's OA Data Science Journal soon to be announced!

DC1 Data Citation Principles



Global Roads Data

Review of Global Roads Data Development Methodologies: http://bit.ly/globalroads-methods

Capacity Building

CODATA Training in Big Data Science
Beijing, 4-20 June 2014
http://bit.ly/CODATA-China_Training_2104-Call

Training Workshop on Open Datafor Science and Sustainability in Developing Countries
Kenya, Jomo Kenyatta University of Science and Technology
http://bit.ly/codata-training-jkuat

CODATA

International Council for Science : Committee on Data for Science and Technology

CODATA Task Groups 2012 - 2014

Through its Task Groups, CODATA executes an ambitious international scientific agenda, addressing major data needs and policy issues in a broad range of subjects. These activities are selected at the biennial CODATA General Assemblies.

The following were approved or renewed at the October 2012 General Assembly in Taipei:

- Advancing Informatics for Microbiology
- Anthropometric Data and Engineering
- Data at Risk
- Data Citation Standards and Practices
- Earth and Space Science Data Interoperability
- Exchangeable Materials Data Representation to Support Scientific Research and Education
- Fundamental Physical Constants
- Global Information Commons for Science Initiative
- Linked Open Data for Global Disaster Risk Research
- Octopus: Mining Space and Terrestrial Data for Improved Weather, Climate and Agriculture Predictions
- Global Roads Data Development
- Preservation of and Access to Scientific and Technical Data in/for/with Developing Countries (PASTD)

Tweets Follow

CODATA @CODATANews 9h

2014 #CODATA General Assembly will be held on 6-7 Nov 2014 at Indian National Science Academy, New Delhi: codata.org/news/6/62/CODA...

Expand

Sarah Jones @sjDCC 23 Jul

RT @kevingashley: I hear rumours that IJDC Volume 9.1 is unleashed, with editorial from #ukdcc's Alex Ball bit.ly/1pcWrmd #jiscmrd

Retweeted by CODATA

Expand

CODATA @CODATANews 23 Jul

US NSF Program Solicitation Partnerships for International Research and Education (PIRE) nsf.gov/pubs/2014/nsf1... deadline 21 October #codata

Expand

CODATA @CODATANews 18 Jul

#CODATA Prize - deadline 21 July

CODATA Task Groups

Fundamental Constants

REVIEWS OF MODERN PHYSICS, VOLUME 84, OCTOBER–DECEMBER 2012

CODATA recommended values of the fundamental physical constants: 2010*

Peter J. Mohr,† Barry N. Taylor,‡ and David B. Newell§
National Institute of Standards and Technology, Gaithersburg, Maryland 20899-8420, USA

(published 13 November 2012)

This paper gives the 2010 self-consistent set of values of the basic constants and conversion factors of physics and chemistry recommended by the Committee on Data for Science and Technology (CODATA) for international use. The 2010 adjustment takes into account the data considered in the 2006 adjustment as well as the data that became available from 1 January 2007, after the closing date of that adjustment, until 31 December 2010, the closing date of the new adjustment. Further, it describes in detail the adjustment of the values of the constants, including the selection of the final set of input data based on the results of least-squares analyses. The 2010 set replaces the previously recommended 2006 CODATA set and may also be found on the World Wide Web at physics.nist.gov/constants.

DOI: 10.1103/RevModPhys.84.1527 PACS numbers: 06.20.Jr, 12.20.−m

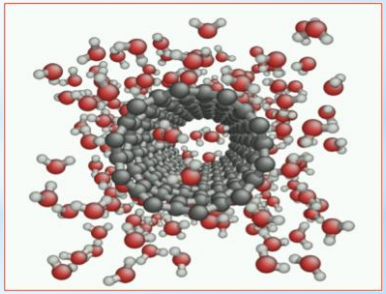
*Since 1969
*Based on all relevant data available

Data at Risk

Memory of the World

Developing 'Principle Guidelines for Data at Risk': http://bit.ly/DAR-guidelines

CODATA WG on Description of Nanomaterials



CODATA-ICSU Workshop: http://www.codata.info/Nanomaterials/Index-agenda-Nanomaterial.html

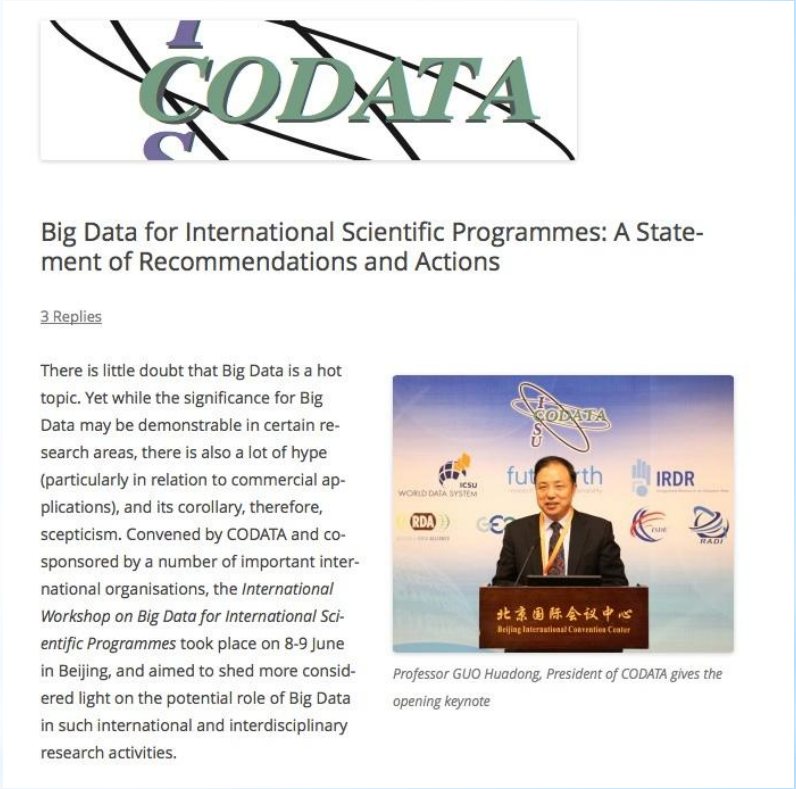
CODATA WG on the Description of Nanomateririals: http://www.codata.org/nanomaterials

Draft Uniform Description System: http://www.codata.org/uploads/Nanomaterials_Framework_for_Uniform_Description_System_2013-10-03.pdf

Future Nano Needs Project: http://www.futurenanoneeds.eu/

CODATA Big Data Workshop Beijing, 8-9 June 2014

- Aim and Objectives
  - to create an international and cross-disciplinary **awareness** of the wide and diverse activities underway in the science-related Big Data sphere;
  - to establish and foster ongoing **coordination and collaboration** for international, interdisciplinary research programmes, such as IRDR and Future Earth, on Big Data issues;
  - to identify ways in which a **CODATA Working Group on Big Data for International Science** programmes can contribute.
- Outcomes
  - **Statement of Recommendations on Big Data for International Sciences Programmes.**
  - White Paper and Editorial
  - CODATA Working Group on Big Data



Big Data for International Scientific Programmes: A Statement of Recommendations and Actions

3 Replies

There is little doubt that Big Data is a hot topic. Yet while the significance for Big Data may be demonstrable in certain research areas, there is also a lot of hype (particularly in relation to commercial applications), and its corollary, therefore, scepticism. Convened by CODATA and co-sponsored by a number of important international organisations, the *International Workshop on Big Data for International Scientific Programmes* took place on 8-9 June in Beijing, and aimed to shed more considered light on the potential role of Big Data in such international and interdisciplinary research activities.

Professor GUO Huadong, President of CODATA gives the opening keynote

http://codata.org/blog/2014/06/16/big-data-for-international-scientific-programmes-a-statement-of-recommendations-and-actions/

Collaboration with RDA

- Memorandum of Understanding and reciprocal membership (with WDS).
- Collaboration in RDA Working Groups/Interest Groups.
- Three co-branded WGs/IGs: Materials, Legal Interoperability and Data Science Summer Schools.
- Co-chair of
  - RDA/WDS IG on Cost Recovery for Data Centers.
  - RDA WG on BioSharing (platform for information on policies, standards and databases).
- Participated on panel on data policy in RDA 3rd Plenary Dublin and on a panel for the Organizational Advisory Board at RDA 4th Plenary Amsterdam
- Proposal to host RDA/CODATA Summer Schools in Data Science and Cloud Computing in the Developing World

RDA
RESEARCH DATA ALLIANCE
Research Data Sharing without barriers

Collaboration with WDS

- Collaboration most obviously in the SciDataCon conference that has just completed.

  - Planning SciDataCon 2016

- Aimed to place data access and reuse issues in the context of major research challenges of global importance.
- Considerable profile on Future Earth Research Themes.
- Plan to collocate Executive Committee and Scientific Committee meetings, with one overlapping day, if possible.
- Ongoing collaboration within the Data Publishing Group.

CODATA-ICSTI Task Group Data Citation, Standards and Practices

For Attribution
Workshop and Report:
http://bit.ly/for_attribution

Out of Cite,Out of Mind
http://bit.ly/out_of_cite

Joint Declaration of Data Citation Principles:
https://www.force11.org/datacitation

Background and Developments:
http://bit.ly/data_citation_principles

DC1
Data Citation Principles

Thanks for your attention!

CODATA Website: http://www.codata.org/
CODATA Blog: http://www.codata.org/blog/
SciDataCon 2014: http://www.scidatacon2014.org/
CODATA General Assembly 2014:
http://www.codata.org/general-assembly

Simon Hodson
Executive Director CODATA
www.codata.org/blog
Email: execdir@codata.org
Twitter: @simonhodson99
Tel (Office): +33 1 45 25 04 96 | Tel (Cell): +33 6 86 30 42 59
CODATA (ICSU Committee on Data for Science and Technology), 5 rue Auguste Vacquerie, 75016 Paris, FRANCE